
Review

A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb

Rosalie Allison, BSc, MSc; Catherine Hayes, BSc; Clodna A M McNulty, MBBS, FRCP; Vicki Young, BSc, PhD
Public Health England, Gloucester, United Kingdom

Corresponding Author:

Rosalie Allison, BSc, MSc
Public Health England
Primary Care and Interventions Unit
Gloucester, GL1 1DQ
United Kingdom
Phone: 44 0208 495 3258
Email: rosie.allison@phe.gov.uk

Abstract

Background: Attention is turning toward increasing the quality of websites and quality evaluation to attract new users and retain existing users.

Objective: This scoping study aimed to review and define existing worldwide methodologies and techniques to evaluate websites and provide a framework of appropriate website attributes that could be applied to any future website evaluations.

Methods: We systematically searched electronic databases and gray literature for studies of website evaluation. The results were exported to EndNote software, duplicates were removed, and eligible studies were identified. The results have been presented in narrative form.

Results: A total of 69 studies met the inclusion criteria. The extracted data included type of website, aim or purpose of the study, study populations (users and experts), sample size, setting (controlled environment and remotely assessed), website attributes evaluated, process of methodology, and process of analysis. Methods of evaluation varied and included questionnaires, observed website browsing, interviews or focus groups, and Web usage analysis. Evaluations using both users and experts and controlled and remote settings are represented. Website attributes that were examined included usability or ease of use, content, design criteria, functionality, appearance, interactivity, satisfaction, and loyalty. Website evaluation methods should be tailored to the needs of specific websites and individual aims of evaluations. GoodWeb, a website evaluation guide, has been presented with a case scenario.

Conclusions: This scoping study supports the open debate of defining the quality of websites, and there are numerous approaches and models to evaluate it. However, as this study provides a framework of the existing literature of website evaluation, it presents a guide of options for evaluating websites, including which attributes to analyze and options for appropriate methods.

(*JMIR Form Res* 2019;3(4):e14372) doi: [10.2196/14372](https://doi.org/10.2196/14372)

KEYWORDS

user experience; usability; human-computer interaction; software testing; quality testing; scoping study

Introduction

Background

Since its conception in the early 1990s, there has been an explosion in the use of the internet, with websites taking a central role in diverse fields such as finance, education, medicine, industry, and business. Organizations are increasingly attempting to exploit the benefits of the World Wide Web and its features as an interface for internet-enabled businesses, information provision, and promotional activities [1,2]. As the

environment becomes more competitive and websites become more sophisticated, attention is turning toward increasing the quality of the website itself and quality evaluation to attract new and retain existing users [3,4]. What determines website quality has not been conclusively established, and there are many different definitions and meanings of the term quality, mainly in relation to the website's purpose [5]. Traditionally, website evaluations have focused on usability, defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [6]." The design of websites and

users' needs go beyond pure usability, as increased engagement and pleasure experienced during interactions with websites can be more important predictors of website preference than usability [7-10]. Therefore, in the last decade, website evaluations have shifted their focus to users' experience, employing various assessment techniques [11], with no universally accepted method or procedure for website evaluation.

Objectives

This scoping study aimed to review and define existing worldwide methodologies and techniques to evaluate websites and provide a simple framework of appropriate website attributes, which could be applied to future website evaluations.

A scoping study is similar to a systematic review as it collects and reviews content in a field of interest. However, scoping studies cover a broader question and do not rigorously evaluate the quality of the studies included [12]. Scoping studies are commonly used in the fields of public services such as health and education, as they are more rapid to perform and less costly in terms of staff costs [13]. Scoping studies can be precursors to a systematic review or stand-alone studies to examine the range of research around a particular topic.

The following research question is based on the need to gain knowledge and insight from worldwide website evaluation to inform the future study design of website evaluations: what website evaluation methodologies can be robustly used to assess users' experience?

To show how the framework of attributes and methods can be applied to evaluating a website, e-Bug, an international educational health website, will be used as a case scenario [14].

Methods

This scoping study followed a 5-stage framework and methodology, as outlined by Arksey and O'Malley [12], involving the following: (1) identifying the research question, as above; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting the results.

Identifying Relevant Studies

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [15], studies for consideration in the review were located by searching the following electronic databases: Excerpta Medica dataBASE, PsycINFO, Cochrane, Cumulative Index to Nursing and Allied Health Literature, Scopus, ACM digital library, and IEEE Xplore SPORTDiscus. The keywords used referred to the following:

- Population: websites
- Intervention: evaluation methodologies
- Outcome: user's experience.

Table 1 shows the specific search criteria for each database. These keywords were also used to search gray literature for unpublished or working documents to minimize publication bias.

Table 1. Full search strategy used to search each electronic database.

Database	Search criteria	Note	Hits (n)
EMBASE ^a	(web* OR internet OR online) AND (user test* OR heuristic)	Search on the field "Title"	689
PsychINFO	evaluation OR usability OR evaluation method* OR measur* OR eye-track* OR eye track*	Search on the field "Title"	816
Cochrane	OR metric* OR rat* OR rank* OR question* OR survey OR stud* OR thinking aloud OR think aloud OR observ* OR complet* OR evaluat*	Search on the fields "Title, keywords, abstract"	1004
CINAHL ^b	OR attribut* OR task*) AND (satisf* OR quality OR efficien* OR task efficiency OR effective* OR appear* OR content	Search on the field "Title"	263
Scopus	OR loyal* OR promot* OR adequa* OR eas* OR user* OR experien*);	Search on the field "Title"	3714
ACM ^c Digital Library	Publication date=between 2006 and 2016; Language published in=English	Search on the field "Title"	89
IEEE ^d Xplore	(web) AND (evaluat) AND (satisf* OR user* OR quality*); Publication date=between 2006 and 2016; Language published in=English	Search on the field "Title"	82

^aEMBASE: Excerpta Medica database.

^bCINAHL: Cumulative Index to Nursing and Allied Health Literature.

^cACM: Association for Computing Machinery.

^dIEEE: Institute of Electrical and Electronics Engineers.

Study Selection

Once all sources had been systematically searched, the list of citations was exported to EndNote software to identify eligible studies. By scanning the title, and abstract if necessary, studies that did not fit the inclusion criteria were removed by 2 researchers (RA and CH). As abstracts are not always representative of the full study that follows or capture the full scope [16], if the title and abstract did not provide sufficient information, the full manuscript was examined to ascertain whether they met all the inclusion criteria, which included (1) studies focused on websites, (2) studies of evaluative methods (eg, use of questionnaire and task completion), (3) studies that reported outcomes that affect the user's experience (eg, quality, satisfaction, efficiency, effectiveness without necessarily focusing on methodology), (4) studies carried out between 2006 and 2016, (5) studies published in English, and (6) type of study (any study design that is appropriate).

Exclusion criteria included (1) studies that focus on evaluations using solely experts and are not transferrable to user evaluations; (2) studies that are in the form of electronic book or are not freely available on the Web or through OpenAthens, the University of Bath library, or the University of the West of England library; (3) studies that evaluate banking, electronic commerce (e-commerce), or online libraries' websites and do not have transferrable measures to a range of other websites; (4) studies that report exclusively on minority or special needs groups (eg, blind or deaf users); and (5) studies that do not meet all the inclusion criteria.

Charting the Data

The next stage involved *charting* key items of information obtained from studies being reviewed. *Charting* [17] describes a technique for synthesizing and interpreting qualitative data by sifting, charting, and sorting material according to key issues and themes. This is similar to a systematic review in which the

process is called data extraction. The data extracted included general information about the study and specific information relating to, for instance, the study population or target, the type of intervention, outcome measures employed, and the study design.

The information of interest included the following: type of website, aim or purpose of the study, study populations (users and experts), sample size, setting (laboratory, real life, and remotely assessed), website attributes evaluated, process of methodology, and process of analysis.

NVivo version 10.0 software was used for this stage by 2 researchers (RA and CH) to chart the data.

Collating, Summarizing, and Reporting the Results

Although the scoping study does not seek to assess the quality of evidence, it does present an overview of all material reviewed with a narrative account of findings.

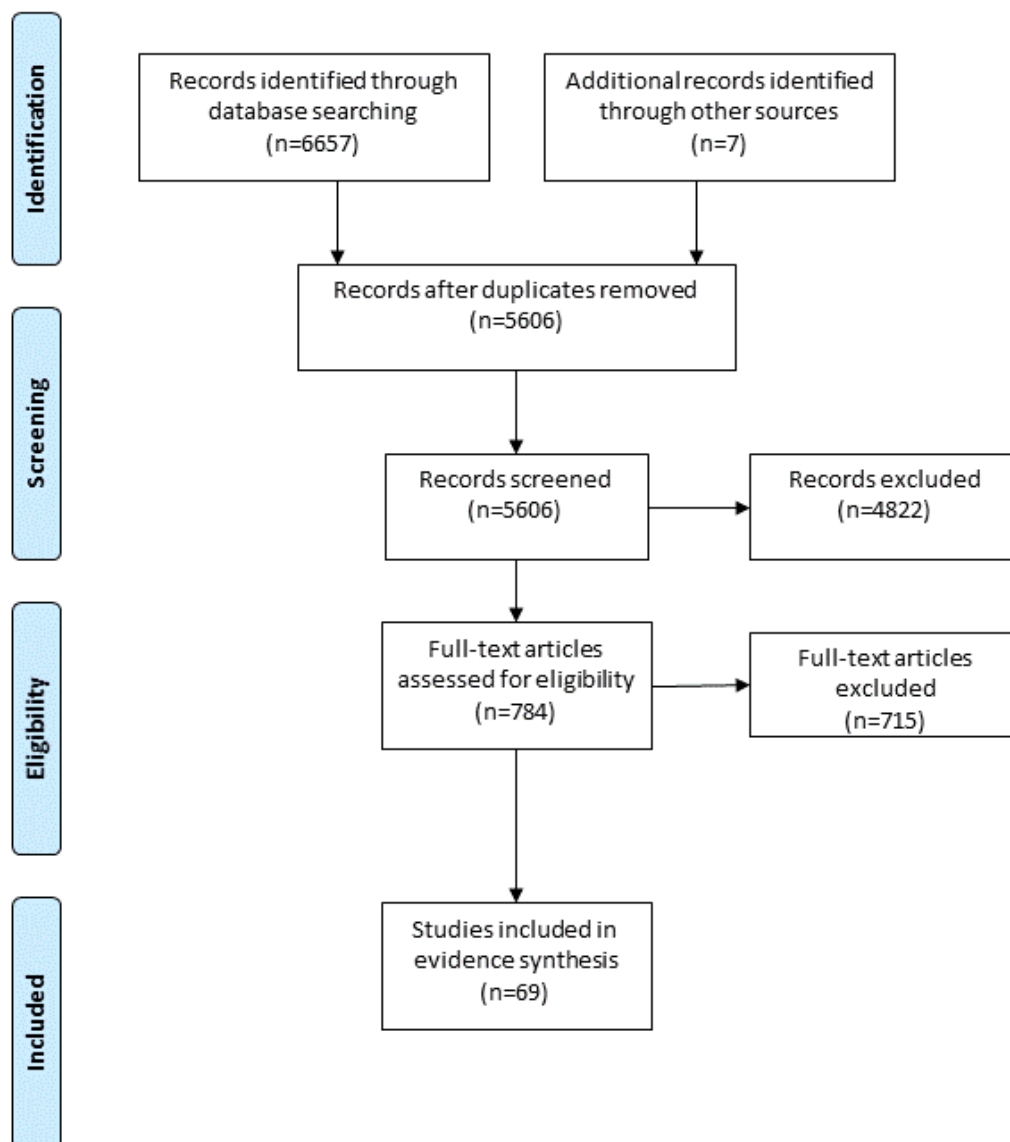
Ethics Approval and Consent to Participate

As no primary research was carried out, no ethical approval was required to undertake this scoping study. No specific reference was made to any of the participants in the individual studies, nor does this study infringe on their rights in any way.

Results

Study Selection

The electronic database searches produced 6657 papers; a further 7 papers were identified through other sources. After removing duplicates (n=1058), 5606 publications remained. After titles and abstracts were examined, 784 full-text papers were read and assessed further for eligibility. Of those, 69 articles were identified as suitable by meeting all the inclusion criteria (Figure 1).

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of search results.

Study Characteristics

Population

Studies referred to or used a mixture of users (72%) and experts (39%) to evaluate their websites; 54% used a controlled environment, and 26% evaluated websites remotely ([Multimedia Appendix 1](#) [2-4,11,18-85]). Remote usability, in its most basic form, involves working with participants who are not in the same physical location as the researcher, employing techniques such as live screen sharing or questionnaires. Advantages to remote website evaluations include the ability to evaluate using a larger number of participants as travel time and costs are not a factor, and participants are able to partake at a time that is appropriate to them, increasing the likelihood of participation and the possibility of a greater diversity of participants [18]. However, the disadvantages of remote website evaluations, in comparison with a controlled setting, are that system performance, network traffic, and the participant's computer setup can all affect the results.

A variety of types of websites evaluated were included in this review including government (9%), online news (6%), education (1%), university (12%), and sports organizations (4%). The aspects of quality considered, and their relative importance varied according to the type of website and the goals to be achieved by the users. For example, criteria such as ease of paying or security are not very important to educational websites, whereas they are especially important for online shopping. In this sense, much attention must be paid when evaluating the quality of a website, establishing a specific context of use and purpose [19].

The context of the participants was also discussed, in relation to the generalizability of results. For example, when evaluations used potential or current users of their website, it was important that computer literacy was reflective of all users [20]. This could mean ensuring that participants with a range of computer abilities and experiences were used so that results were not biased to the most or least experienced users.

Intervention

A total of 43 evaluation methodologies were identified in the 69 studies in this review. Most of them were variations of similar methodologies, and a brief description of each is provided in [Multimedia Appendix 2](#). [Multimedia Appendix 3](#) shows the methods used or described in each study.

Questionnaire

Use of questionnaires was the most common methodology referred to (37/69, 54%), including questions to rank or rate attributes and open questions to allow text feedback and suggested improvements. Questionnaires were used in a combination of before or after usability testing to assess usability and overall user experience.

Observed Browsing the Website

Browsing the website using a form of task completion with the participant, such as cognitive walkthrough, was used in 33/69 studies (48%), whereby an expert evaluator used a detailed procedure to simulate task execution and browse all particular solution paths, examining each action while determining if expected user's goals and memory content would lead to choosing a correct option [30]. Screen capture was often used (n=6) to record participants' navigation through the website, and eye tracking was used (n=7) to assess where the eye focuses on each page or the motion of the eye as an individual views a Web page. The think-aloud protocol was used (n=10) to encourage users to express out loud what they were looking at, thinking, doing, and feeling, as they performed tasks. This allows observers to see and understand the cognitive processes associated with task completion. Recording the time to complete tasks (n=6) and mouse movement or clicks (n=8) were used to assess the efficiency of the websites.

Qualitative Data Collection

Several forms of qualitative data collection were used in 27/69 studies (39%). Observed browsing, interviews, and focus groups were used either before or after the use of the website. Pre-website-use, qualitative research was often used to collect details of which website attributes were important for participants or what weighting participants would give to each attribute. Postevaluation, qualitative techniques were used to collate feedback on the quality of the website and any suggestions for improvements.

Automated Usability Evaluation Software

In 9/69 studies (13%), automated usability evaluation focused on developing software, tools, and techniques to speed evaluation (rapid), tools that reach a wider audience for usability testing (remote), and tools that have built-in analyses features (automated). The latter can involve assessing server logs, website coding, and simulations of user experience to assess usability [42].

Card Sorting

A technique that is often linked with assessing navigability of a website, card sorting, is useful for discovering the logical structure of an unsorted list of statements or ideas by exploring how people group items and structures that maximize the probability of users finding items (5/69 studies, 7%). This can assist with determining effective website structure.

Web Usage Analysis

Of 69 studies, 3 studies used Web usage analysis or Web analytics to identify browsing patterns by analyzing the participants' navigational behavior. This could include tracking at the widget level, that is, combining knowledge of the mouse coordinates with elements such as buttons and links, with the layout of the HTML pages, enabling complete tracking of all user activity.

Outcomes (Attributes Used to Evaluate Websites)

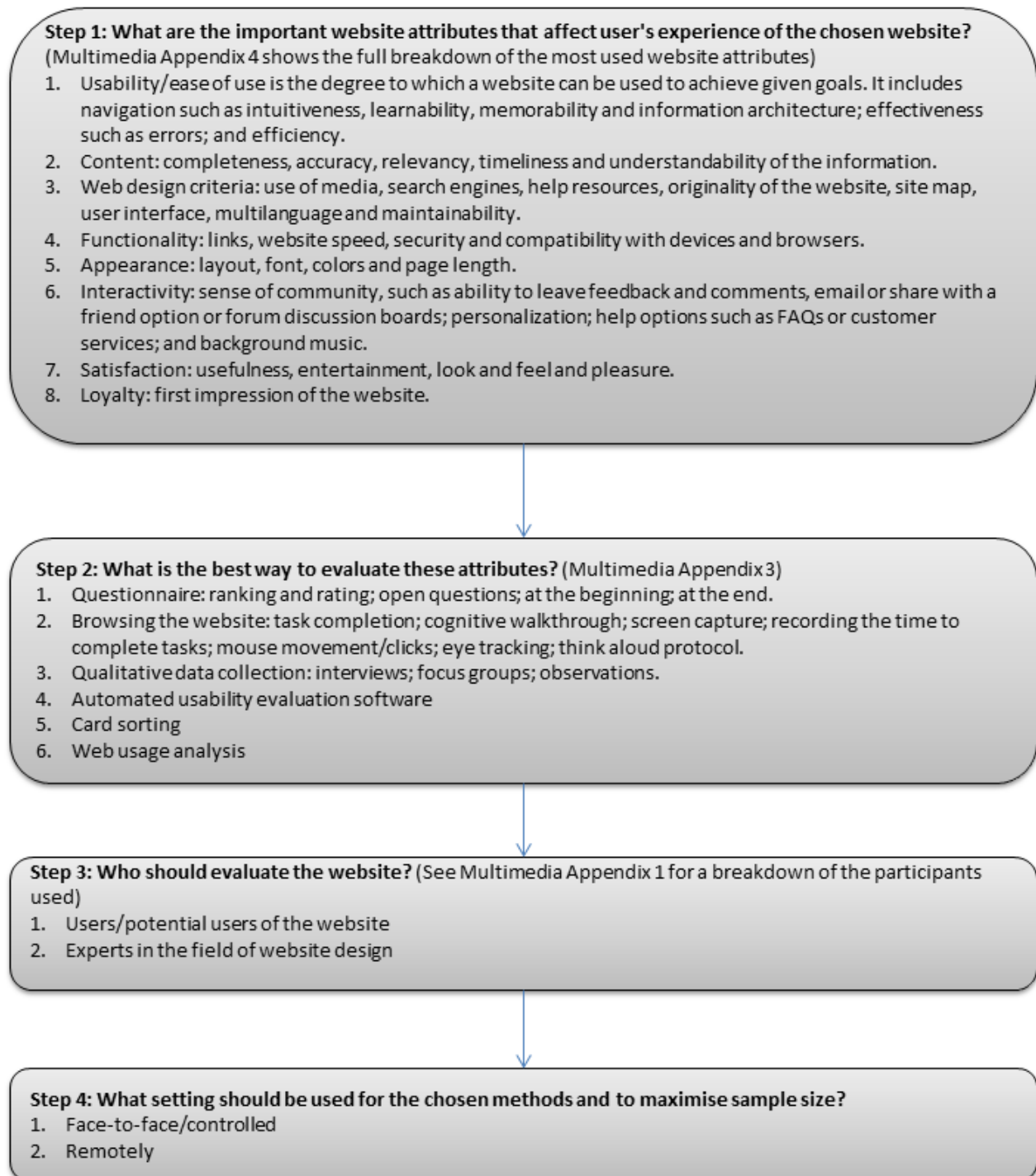
Often, different terminology for website attributes was used to describe the same or similar concepts ([Multimedia Appendix 4](#)). The most used website attributes that were assessed can be broken down into 8 broad categories and further subcategories:

1. Usability or ease of use is the degree to which a website can be used to achieve given goals (n=58). It includes navigation such as intuitiveness, learnability, memorability, and information architecture; effectiveness such as errors; and efficiency.
2. Content (n=41) includes completeness, accuracy, relevancy, timeliness, and understandability of the information.
3. Web design criteria (n=29) include use of media, search engines, help resources, originality of the website, site map, user interface, multilanguage, and maintainability.
4. Functionality (n=31) includes links, website speed, security, and compatibility with devices and browsers.
5. Appearance (n=26) includes layout, font, colors, and page length.
6. Interactivity (n=25) includes sense of community, such as ability to leave feedback and comments and email or share with a friend option or forum discussion boards; personalization; help options such as frequently answered questions or customer services; and background music.
7. Satisfaction (n=26) includes usefulness, entertainment, look and feel, and pleasure.
8. Loyalty (n=8) includes first impression of the website.

Discussion

GoodWeb: Website Evaluation Guide

As there was such a range of methods used, a suggested guide of options for evaluating websites is presented below ([Figure 2](#)), coined GoodWeb, and applied to an evaluation of e-Bug, an international educational health website [14]. Allison et al [86] show the full details of how GoodWeb has been applied and outcomes of the e-Bug website evaluation.

Figure 2. Framework for website evaluation.

Step 1. What Are the Important Website Attributes That Affect User's Experience of the Chosen Website?

Usability or ease of use, content, Web design criteria, functionality, appearance, interactivity, satisfaction, and loyalty were the umbrella terms that encompassed the website attributes identified or evaluated in the 69 studies in this scoping study. [Multimedia Appendix 4](#) contains a summary of the most used website attributes that have been assessed. Recent website evaluations have shifted focus from usability of websites to an overall user's experience of website use. A decision on which website attributes to evaluate for specific websites could come

from interviews or focus groups with users or experts or a literature search of attributes used in similar evaluations.

Application

In the scenario of evaluating e-Bug or similar educational health websites, the attributes chosen to assess could be the following:

1. Appearance: colors, fonts, media or graphics, page length, style consistency, and first impression
2. Content: clarity, completeness, current and timely information, relevance, reliability, and uniqueness
3. Interactivity: sense of community and modern features

4. Ease of use: home page indication, navigation, guidance, and multilanguage support
5. Technical adequacy: compatibility with other devices, load time, valid links, and limited use of special plug-ins
6. Satisfaction: loyalty

These cover the main website attributes appropriate for an educational health website. If the website did not currently have features such as search engines, site map, background music, it may not be appropriate to evaluate these, but may be better suited to question whether they would be suitable additions to the website; or these could be combined under the heading *modern features*. Furthermore, security may not be a necessary attribute to evaluate if participant identifiable information or bank details are not needed to use the website.

Step 2. What Is the Best Way to Evaluate These Attributes?

Often, a combination of methods is suitable to evaluate a website, as 1 method may not be appropriate to assess all attributes of interest [29] (see [Multimedia Appendix 3](#) for a summary of the most used methods for evaluating websites). For example, screen capture of task completion may be appropriate to assess the efficiency of a website but would not be the chosen method to assess loyalty. A questionnaire or qualitative interview may be more appropriate for this attribute.

Application

In the scenario of evaluating e-Bug, a questionnaire before browsing the website would be appropriate to rank the importance of the selected website attributes, chosen in step 1. It would then be appropriate to observe browsing of the website, collecting data on completion of typical task scenarios, using the screen capture function for future reference. This method could be used to evaluate the effectiveness (number of tasks successfully completed), efficiency (whether the most direct route through the website was used to complete the task), and learnability (whether task completion is more efficient or effective second time of trying). It may then be suitable to use a follow-up questionnaire to rate e-Bug against the website attributes previously ranked. The attribute ranking and rating could then be combined to indicate where the website performs well and areas for improvement.

Step 3: Who Should Evaluate the Website?

Both users and experts can be used to evaluate websites. Experts are able to identify areas for improvements, in relation to usability; whereas, users are able to appraise quality as well as identify areas for improvement. In this respect, users are able to fully evaluate user's experience, where experts may not be able to.

Application

For this reason, it may be more appropriate to use current or potential users of the website for the scenario of evaluating e-Bug.

Step 4: What Setting Should Be Used?

A combination of controlled and remote settings can be used, depending on the methods chosen. For example, it may be appropriate to collect data via a questionnaire, remotely, to increase sample size and reach a more diverse audience, whereas a controlled setting may be more appropriate for task completion using eye-tracking methods.

Strengths and Limitations

A scoping study differs from a systematic review, in that it does not critically appraise the quality of the studies before extracting or *charting* the data. Therefore, this study cannot compare the effectiveness of the different methods or methodologies in evaluating the website attributes. However, what it does do is review and summarize a huge amount of literature, from different sources, in a format that is understandable and informative for future designs of website evaluations.

Furthermore, studies that evaluate banking, e-commerce, or online libraries' websites and do not have transferrable measures to a range of other websites were excluded from this study. This decision was made to limit the number of studies that met the remaining inclusion criteria, and it was deemed that the website attributes for these websites would be too specialist and not necessarily transferable to a range of websites. Therefore, the findings of this study may not be generalizable to all types of website. However, [Multimedia Appendix 1](#) shows that data were extracted from a very broad range of websites when it was deemed that the information was transferrable to a range of other websites.

A robust website evaluation can identify areas for improvement to both fulfill the goals and desires of its users [62] and influence their perception of the organization and overall quality of resources [48]. An improved website could attract and retain more online users; therefore, an evidence-based website evaluation guide is essential.

Conclusions

This scoping study emphasizes the fact that the debate about how to define the quality of websites remains open, and there are numerous approaches and models to evaluate it. [Multimedia Appendix 2](#) shows existing methodologies or tools that can be used to evaluate websites. Many of these are variations of similar approaches; therefore, it is not strictly necessary to use these tools at face value; however, some could be used to guide analysis, following data collection. By following steps 1 to 4 of GoodWeb, the framework suggested in this study, taking into account the desired participants and setting and website evaluation methods, can be tailored to the needs of specific websites and individual aims of evaluations.

Acknowledgments

This work was supported by the Primary Care Unit, Public Health England. This study is not applicable as secondary research.

Authors' Contributions

RA wrote the protocol with input from CH, CM, and VY. RA and CH conducted the scoping review. RA wrote the final manuscript with input from CH, CM, and VY. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of included studies, including information on the participant.

[\[PDF File \(Adobe PDF File\), 133 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Interventions: methodologies and tools to evaluate websites.

[\[PDF File \(Adobe PDF File\), 85 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Methods used or described in each study.

[\[XLSX File \(Microsoft Excel File\), 20 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Summary of the most used website attributes evaluated.

[\[XLSX File \(Microsoft Excel File\), 26 KB-Multimedia Appendix 4\]](#)

References

1. Straub DW, Watson RT. Research Commentary: Transformational Issues in Researching IS and Net-Enabled Organizations. *Info Syst Res* 2001;12(4):337-345. [doi: [10.1287/isre.12.4.337.9706](https://doi.org/10.1287/isre.12.4.337.9706)]
2. Bairamzadeh S, Bolhari A. Investigating factors affecting students' satisfaction of university websites. In: 2010 3rd International Conference on Computer Science and Information Technology. 2010 Presented at: ICCSIT'10; July 9-11, 2010; Chengdu, China p. 469-473. [doi: [10.1109/iccst.2010.5565050](https://doi.org/10.1109/iccst.2010.5565050)]
3. Fink D, Nyaga C. Evaluating web site quality: the value of a multi paradigm approach. *Benchmarking* 2009;16(2):259-273. [doi: [10.1108/14635770910948259](https://doi.org/10.1108/14635770910948259)]
4. Markaki OI, Charilas DE, Askounis D. Application of Fuzzy Analytic Hierarchy Process to Evaluate the Quality of E-Government Web Sites. In: Proceedings of the 2010 Developments in E-systems Engineering. 2010 Presented at: DeSE'10; September 6-8, 2010; London, UK p. 219-224. [doi: [10.1109/dese.2010.42](https://doi.org/10.1109/dese.2010.42)]
5. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *J Am Med Assoc* 2002;287(20):2691-2700. [doi: [10.1001/jama.287.20.2691](https://doi.org/10.1001/jama.287.20.2691)] [Medline: [12020305](https://pubmed.ncbi.nlm.nih.gov/12020305/)]
6. International Organization for Standardization. ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability. Switzerland: International Organization for Standardization; 1998.
7. Hartmann J, Sutcliffe A, Angeli AD. Towards a theory of user judgment of aesthetics and user interface quality. *ACM Trans Comput-Hum Interact* 2008;15(4):1-30. [doi: [10.1145/1460355.1460357](https://doi.org/10.1145/1460355.1460357)]
8. Bargas-Avila JA, Hornbæk K. Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI'11; May 7-12, 2011; Vancouver, BC, Canada p. 2689-2698. [doi: [10.1145/1978942.1979336](https://doi.org/10.1145/1978942.1979336)]
9. Hassenzahl M, Tractinsky N. User experience - a research agenda. *Behav Info Technol* 2006;25(2):91-97. [doi: [10.1080/01449290500330331](https://doi.org/10.1080/01449290500330331)]
10. Aranyi G, van Schaik P. Testing a model of user-experience with news websites. *J Assoc Soc Inf Sci Technol* 2016;67(7):1555-1575. [doi: [10.1002/asi.23462](https://doi.org/10.1002/asi.23462)]
11. Tsai W, Chou W, Lai C. An effective evaluation model and improvement analysis for national park websites: a case study of Taiwan. *Tour Manag* 2010;31(6):936-952. [doi: [10.1016/j.tourman.2010.01.016](https://doi.org/10.1016/j.tourman.2010.01.016)]
12. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Method* 2005;8(1):19-32. [doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616)]
13. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Res Policy Syst* 2008;6:7 [FREE Full text] [doi: [10.1186/1478-4505-6-7](https://doi.org/10.1186/1478-4505-6-7)] [Medline: [18613961](https://pubmed.ncbi.nlm.nih.gov/18613961/)]

14. e-Bug. 2018. Welcome to the e-Bug Teachers Area! URL:https://e-bug.eu/eng_home.aspx?cc=eng&ss=1&t=Welcome%20to%20e-Bug [accessed 2019-08-23]
15. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: [10.1371/journal.pmed.1000100](https://doi.org/10.1371/journal.pmed.1000100)] [Medline: [19621070](https://pubmed.ncbi.nlm.nih.gov/19621070/)]
16. Badger D, Nursten J, Williams P, Woodward M. Should All Literature Reviews be Systematic? *Eval Res Educ* 2000;14(3-4):220-230. [doi: [10.1080/09500790008666974](https://doi.org/10.1080/09500790008666974)]
17. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess B, editors. *Analyzing Qualitative Data*. Abingdon-on-Thames: Routledge; 2002:187-208.
18. Thomsett-Scott BC. Web site usability with remote users: Formal usability studies and focus groups. *J Libr Adm* 2006;45(3-4):517-547. [doi: [10.1300/j111v45n03_14](https://doi.org/10.1300/j111v45n03_14)]
19. Moreno JM, Morales del Castillo JM, Porcel C, Herrera-Viedma E. A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach. *Soft Comput* 2010;14(8):887-897. [doi: [10.1007/s00500-009-0472-7](https://doi.org/10.1007/s00500-009-0472-7)]
20. Alva M, Martínez A, Labra Gayo J, Del Carmen Suárez M, Cueva J, Sagástegui H. Proposal of a tool of support to the evaluation of user in educative web sites. 2008 Presented at: 1st World Summit on the Knowledge Society, WSKS 2008; 2008; Athens p. 149-157. [doi: [10.1007/978-3-540-87781-3_17](https://doi.org/10.1007/978-3-540-87781-3_17)]
21. Usability: Home. Usability Evaluation Basics URL:<https://www.usability.gov/what-and-why/usability-evaluation.html> [accessed 2019-08-24]
22. AddThis: Get more likes, shares and follows with smart. 10 Criteria for Better Website Usability: Heuristics Cheat Sheet URL:<http://www.addthis.com/blog/2015/02/17/10-criteria-for-better-website-usability-heuristics-cheat-sheet/#.V712QfkrJD8> [accessed 2019-08-24]
23. Akgül Y. Quality evaluation of E-government websites of Turkey. In: Proceedings of the 2016 11th Iberian Conference on Information Systems and Technologies. 2016 Presented at: CISTI'16; June 15-18 2016; Las Palmas, Spain p. 1-7. [doi: [10.1109/cisti.2016.7521567](https://doi.org/10.1109/cisti.2016.7521567)]
24. Al Zaghoul FA, Al Nsour AJ, Rababah OM. Ranking Quality Factors for Measuring Web Service Quality. In: Proceedings of the 1st International Conference on Intelligent Semantic Web-Services and Applications. 2010 Presented at: 1st ACM Jordan Professional Chapter ISWSA Annual - International Conference on Intelligent Semantic Web-Services and Applications, ISWSA'10; June 14-16, 2010; Amman, Jordan. [doi: [10.1145/1874590.1874608](https://doi.org/10.1145/1874590.1874608)]
25. Alharbi A, Mayhew P. Users' Performance in Lab and Non-lab Environments Through Online Usability Testing: A Case of Evaluating the Usability of Digital Academic Libraries' Websites. In: 2015 Science and Information Conference. 2015 Presented at: SAI'15; July 28-30, 2015; London, UK p. 151-161. [doi: [10.1109/sai.2015.7237139](https://doi.org/10.1109/sai.2015.7237139)]
26. Aliyu M, Mahmud M, Md Tap AO. Preliminary Investigation of Islamic Websites Design & Content Feature: A Heuristic Evaluation From User Perspective. In: Proceedings of the 2010 International Conference on User Science and Engineering. 2010 Presented at: iUser'10; December 13-15, 2010; Shah Alam, Malaysia p. 262-267. [doi: [10.1109/iuser.2010.5716763](https://doi.org/10.1109/iuser.2010.5716763)]
27. Aliyu M, Mahmud M, Tap AO, Nassr RM. Evaluating Design Features of Islamic Websites: A Muslim User Perception. In: Proceedings of the 2013 5th International Conference on Information and Communication Technology for the Muslim World. 2013 Presented at: ICT4M'13; March 26-27, 2013; Rabat, Morocco. [doi: [10.1109/ict4m.2013.6518878](https://doi.org/10.1109/ict4m.2013.6518878)]
28. Al-Radaideh QA, Abu-Shanab E, Hamam S, Abu-Salem H. Usability evaluation of online news websites: a user perspective approach. *World Acad Sci Eng Technol* 2011;74:1058-1066 [FREE Full text]
29. Aranyi G, van Schaik P, Barker P. Using think-aloud and psychometrics to explore users' experience with a news web site. *Interact Comput* 2012;24(2):69-77. [doi: [10.1016/j.intcom.2012.01.001](https://doi.org/10.1016/j.intcom.2012.01.001)]
30. Arrue M, Fajardo I, Lopez JM, Vigo M. Interdependence between technical web accessibility and usability: its influence on web quality models. *Int J Web Eng Technol* 2007;3(3):307-328. [doi: [10.1504/ijwet.2007.012059](https://doi.org/10.1504/ijwet.2007.012059)]
31. Arrue M, Vigo M, Abascal J. Quantitative metrics for web accessibility evaluation. 2005 Presented at: Proceedings of the ICWE 2005 Workshop on Web Metrics and Measurement; 2005; Sydney.
32. Atterer R, Wnuk M, Schmidt A. Knowing the User's Every Move: User Activity Tracking for Website Usability Evaluation and Implicit Interaction. In: Proceedings of the 15th international conference on World Wide. 2006 Presented at: WWW'06; May 23-26, 2006; Edinburgh, Scotland p. 203-212. [doi: [10.1145/1135777.1135811](https://doi.org/10.1145/1135777.1135811)]
33. Bahry F, Masrom M, Masrek M. Website evaluation measures, website user engagement and website credibility for municipal website. *ARNP J Eng Appl Sci* 2015;10(23):18228-18238. [doi: [10.1109/iuser.2016.7857963](https://doi.org/10.1109/iuser.2016.7857963)]
34. Bañón-Gomis A, Tomás-Miquel JV, Expósito-Langa M. Improving user experience: a methodology proposal for web usability measurement. In: *Strategies in E-Business: Positioning and Social Networking in Online Markets*. New York City: Springer US; 2014:123-145.
35. Barnes SJ, Vidgen RT. Data triangulation and web quality metrics: a case study in e-government. *Inform Manag* 2006;43(6):767-777. [doi: [10.1016/j.im.2006.06.001](https://doi.org/10.1016/j.im.2006.06.001)]
36. Bolchini D, Garzotto F. Quality of Web Usability Evaluation Methods: An Empirical Study on MiLE+. In: Proceedings of the 2007 international conference on Web information systems engineering. 2007 Presented at: WISE'07; December 3-3, 2007; Nancy, France p. 481-492. [doi: [10.1007/978-3-540-77010-7_47](https://doi.org/10.1007/978-3-540-77010-7_47)]

37. Chen FH, Tzeng G, Chang CC. Evaluating the enhancement of corporate social responsibility websites quality based on a new hybrid MADM model. *Int J Inf Technol Decis Mak* 2015;14(03):697-724. [doi: [10.1142/s0219622015500121](https://doi.org/10.1142/s0219622015500121)]
38. Cherfi SS, Tuan AD, Comyn-Wattiau I. An Exploratory Study on Websites Quality Assessment. In: Proceedings of the 32nd International Conference on Conceptual Modeling Workshops. 2013 Presented at: ER'13; November 11-13, 2014; Hong Kong, China p. 170-179. [doi: [10.1007/978-3-319-14139-8_19](https://doi.org/10.1007/978-3-319-14139-8_19)]
39. Chou W, Cheng Y. A hybrid fuzzy MCDM approach for evaluating website quality of professional accounting firms. *Expert Sys Appl* 2012;39(3):2783-2793. [doi: [10.1016/j.eswa.2011.08.138](https://doi.org/10.1016/j.eswa.2011.08.138)]
40. Churm T. Usability Geek. 2012 Jul 9. An Introduction To Website Usability Testing URL:<http://usabilitygeek.com/an-introduction-to-website-usability-testing/> [accessed 2019-08-24]
41. Demir Y, Gozum S. Evaluation of quality, content, and use of the web site prepared for family members giving care to stroke patients. *Comput Inform Nurs* 2015 Sep;33(9):396-403. [doi: [10.1097/CIN.000000000000165](https://doi.org/10.1097/CIN.000000000000165)] [Medline: [26125455](https://pubmed.ncbi.nlm.nih.gov/26125455/)]
42. Dominic P, Jati H, Hanim S. University website quality comparison by using non-parametric statistical test: a case study from Malaysia. *Int J Oper Res* 2013;16(3):349-374. [doi: [10.1504/ijor.2013.052335](https://doi.org/10.1504/ijor.2013.052335)]
43. Elling S, Lentz L, de Jong M, van den Bergh H. Measuring the quality of governmental websites in a controlled versus an online setting with the 'Website Evaluation Questionnaire'. *Gov Inf Q* 2012;29(3):383-393. [doi: [10.1016/j.giq.2011.11.004](https://doi.org/10.1016/j.giq.2011.11.004)]
44. Fang X, Holsapple CW. Impacts of navigation structure, task complexity, and users' domain knowledge on web site usability—an empirical study. *Inf Syst Front* 2011;13(4):453-469. [doi: [10.1007/s10796-010-9227-3](https://doi.org/10.1007/s10796-010-9227-3)]
45. Fernandez A, Abrahão S, Insfran E. A systematic review on the effectiveness of web usability evaluation methods. In: Proceedings of the 16th International Conference on Evaluation & Assessment in Software Engineering. 2012 Presented at: EASE'12; May 14-15, 2012; Ciudad Real, Spain. [doi: [10.1049/ic.2012.0007](https://doi.org/10.1049/ic.2012.0007)]
46. Flavián C, Guinalú M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Inform Manag* 2006;43(1):1-14. [doi: [10.1016/j.im.2005.01.002](https://doi.org/10.1016/j.im.2005.01.002)]
47. Flavián C, Guinalú M, Gurrea R. The influence of familiarity and usability on loyalty to online journalistic services: the role of user experience. *J Retail Consum Serv* 2006;13(5):363-375. [doi: [10.1016/j.jretconser.2005.11.003](https://doi.org/10.1016/j.jretconser.2005.11.003)]
48. Gonzalez ME, Quesada G, Davis J, Mora-Monge C. Application of quality management tools in the evaluation of websites: the case of sports organizations. *Qual Manage J* 2015;22(1):30-46. [doi: [10.1080/10686967.2015.11918417](https://doi.org/10.1080/10686967.2015.11918417)]
49. Harrison C, Pétrie H. Deconstructing Web Experience: More Than Just Usability and Good Design. In: Proceedings of the 12th international conference on Human-computer interaction: applications and services. 2007 Presented at: HCI'07; July 22-27, 2007; Beijing, China p. 889-898. [doi: [10.1007/978-3-540-73111-5_98](https://doi.org/10.1007/978-3-540-73111-5_98)]
50. Hart D, Portwood DM. Usability Testing of Web Sites Designed for Communities of Practice: Tests of the IEEE Professional Communication Society (PCS) Web Site Combining Specialized Heuristic Evaluation and Task-based User Testing. In: Proceedings of the 2009 IEEE International Professional Communication Conference. 2009 Presented at: 2009 IEEE International Professional Communication Conference; July 19-22, 2009; Waikiki, HI, USA. [doi: [10.1109/ipcc.2009.5428218](https://doi.org/10.1109/ipcc.2009.5428218)]
51. Hedegaard S, Simonsen JG. Extracting Usability and User Experience Information From Online User Reviews. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI'13; April 27-May 2, 2013; Paris, France p. 2089-2098. [doi: [10.1145/2470654.2481286](https://doi.org/10.1145/2470654.2481286)]
52. Herrera F, Herrera-Viedma E, Martínez L, Ferez LG, López-Herrera AG, Alonso S. A multi-granular linguistic hierarchical model to evaluate the quality of web site services. In: Mathew S, Mordeson JN, Malik DS, editors. *Studies in Fuzziness and Soft Computing*. New York City: Springer; 2006:247-274.
53. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. *Health Promot J Austr* 2008 Apr;19(1):29-35. [doi: [10.1071/HE08029](https://doi.org/10.1071/HE08029)] [Medline: [18481929](https://pubmed.ncbi.nlm.nih.gov/18481929/)]
54. Ijaz T, Andlib F. Impact of Usability on Non-technical Users: Usability Testing Through Websites. In: Proceedings of the 2014 National Software Engineering Conference. 2014 Presented at: 2014 National Software Engineering Conference, NSEC 2014; November 11-12, 2014; Rawalpindi, Pakistan. [doi: [10.1109/nsec.2014.6998238](https://doi.org/10.1109/nsec.2014.6998238)]
55. Janiak E, Rhodes E, Foster AM. Translating access into utilization: lessons from the design and evaluation of a health insurance Web site to promote reproductive health care for young women in Massachusetts. *Contraception* 2013 Dec;88(6):684-690. [doi: [10.1016/j.contraception.2013.09.004](https://doi.org/10.1016/j.contraception.2013.09.004)] [Medline: [24120249](https://pubmed.ncbi.nlm.nih.gov/24120249/)]
56. Kaya T. Multi-attribute evaluation of website quality in e-business using an integrated fuzzy AHP-TOPSIS methodology. *Int J Comput Intell Syst* 2010;3(3):301-314. [doi: [10.2991/ijcis.2010.3.3.6](https://doi.org/10.2991/ijcis.2010.3.3.6)]
57. Kincl T, Štrach P. Measuring website quality: asymmetric effect of user satisfaction. *Behav Inform Technol* 2012;31(7):647-657. [doi: [10.1080/0144929x.2010.526150](https://doi.org/10.1080/0144929x.2010.526150)]
58. Koutsabasis P, Istikopoulou TG. Perceived website aesthetics by users and designers: implications for evaluation practice. *Int J Technol Human Interact* 2014;10(2):21-34. [doi: [10.4018/ijthi.2014040102](https://doi.org/10.4018/ijthi.2014040102)]
59. Leuthold S, Schmutz P, Bargas-Avila JA, Tuch AN, Opwis K. Vertical versus dynamic menus on the world wide web: eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Comput Human Behav* 2011;27(1):459-472. [doi: [10.1016/j.chb.2010.09.009](https://doi.org/10.1016/j.chb.2010.09.009)]

60. Longstreet P. Evaluating Website Quality: Applying Cue Utilization Theory to WebQual. In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences. 2010 Presented at: HICSS'10; January 5-8, 2010; Honolulu, HI, USA. [doi: [10.1109/hicss.2010.191](https://doi.org/10.1109/hicss.2010.191)]
61. Manzoor M. Measuring user experience of usability tool, designed for higher educational websites. Middle East J Sci Res 2013;14(3):347-353. [doi: [10.5829/idosi.mejsr.2013.14.3.2053](https://doi.org/10.5829/idosi.mejsr.2013.14.3.2053)]
62. Mashable India. 22 Essential Tools for Testing Your Website's Usability URL:<http://mashable.com/2011/09/30/website-usability-tools/#cNv8ckxZsmqw> [accessed 2019-08-16]
63. Matera M, Rizzo F, Carughi GT. Web usability: principles and evaluation methods. In: Web Engineering. New York City: Springer; 2006:143-180.
64. McClellan MA, Karumur RP, Vogel RI, Petzel SV, Cragg J, Chan D, et al. Designing an educational website to improve quality of supportive oncology care for women with ovarian cancer: an expert usability review and analysis. Int J Hum Comput Interact 2016;32(4):297-307 [FREE Full text] [doi: [10.1080/10447318.2016.1140528](https://doi.org/10.1080/10447318.2016.1140528)] [Medline: [27110082](https://pubmed.ncbi.nlm.nih.gov/27110082/)]
65. Nakamichi N, Shima K, Sakai M, Matsumoto KI. Detecting Low Usability Web Pages Using Quantitative Data of Users' Behavior. In: Proceedings of the 28th international conference on Software engineering. 2006 Presented at: ICSE'06; May 20-28, 2006; Shanghai, China p. 569-576. [doi: [10.1145/1134285.1134365](https://doi.org/10.1145/1134285.1134365)]
66. Nathan RJ, Yeow PH. An empirical study of factors affecting the perceived usability of websites for student internet users. Univ Access Inf Soc 2009;8(3):165-184. [doi: [10.1007/s10209-008-0138-8](https://doi.org/10.1007/s10209-008-0138-8)]
67. Oliver H, Diallo G, de Quincey E, Alexopoulou D, Habermann B, Kostkova P, et al. A user-centred evaluation framework for the Sealife semantic web browsers. BMC Bioinform 2009;10(S10). [doi: [10.1186/1471-2105-10-s10-s14](https://doi.org/10.1186/1471-2105-10-s10-s14)]
68. Paul A, Yadamsuren B, Erdelez S. An Experience With Measuring Multi-User Online Task Performance. In: Proceedings of the 2012 World Congress on Information and Communication Technologies. 2012 Presented at: WICT'12; October 30-November 2, 2012; Trivandrum, India p. 639-644. [doi: [10.1109/wict.2012.6409154](https://doi.org/10.1109/wict.2012.6409154)]
69. Petrie H, Power C. What Do Users Really Care About?: A Comparison of Usability Problems Found by Users and Experts on Highly Interactive Websites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI'12; May 5-10, 2012; Austin, Texas, USA p. 2107-2116. [doi: [10.1145/2207676.2208363](https://doi.org/10.1145/2207676.2208363)]
70. Rekik R, Kallel I. Fuzzy Reduced Method for Evaluating the Quality of Institutional Web Sites. In: Proceedings of the 2011 7th International Conference on Next Generation Web Services Practices. 2011 Presented at: NWeSP'11; October 19-21, 2011; Salamanca, Spain p. 296-301. [doi: [10.1109/nwesp.2011.6088194](https://doi.org/10.1109/nwesp.2011.6088194)]
71. Reynolds E. The secret to patron-centered Web design: cheap, easy, and powerful usability techniques. Comput Librar 2008;28(6):44-47 [FREE Full text]
72. Sheng H, Lockwood NS, Dahal S. Eyes Don't Lie: Understanding Users' First Impressions on Websites Using Eye Tracking. In: Proceedings of the 15th International Conference on Human Interface and the Management of Information: Information and Interaction Design. 2013 Presented at: HCI'13; July 21-26, 2013; Las Vegas, NV, USA p. 635-641. [doi: [10.1007/978-3-642-39209-2_71](https://doi.org/10.1007/978-3-642-39209-2_71)]
73. Swaid SI, Wigand RT. Measuring Web-based Service Quality: The Online Customer Point of View. In: Proceedings of the 13th Americas Conference on Information Systems. 2007 Presented at: AMCIS'07; August 9-12, 2007; Keystone, Colorado, USA p. 778-790.
74. Tan GW, Wei KK. An empirical study of Web browsing behaviour: towards an effective Website design. Elect Commer Res Appl 2006;5(4):261-271. [doi: [10.1016/j.elerap.2006.04.007](https://doi.org/10.1016/j.elerap.2006.04.007)]
75. Tan W, Liu D, Bishu R. Web evaluation: heuristic evaluation vs user testing. Int J Ind Ergonom 2009;39(4):621-627. [doi: [10.1016/j.ergon.2008.02.012](https://doi.org/10.1016/j.ergon.2008.02.012)]
76. Tao D, LeRouge CM, Deckard G, de Leo G. Consumer Perspectives on Quality Attributes in Evaluating Health Websites. In: Proceedings of the 2012 45th Hawaii International Conference on System Sciences. 2012 Presented at: HICSS'12; January 4-7, 2012; Maui, HI, USA. [doi: [10.1109/hicss.2012.180](https://doi.org/10.1109/hicss.2012.180)]
77. The Whole Brain Group. 2011. Conducting a Quick & Dirty Evaluation of Your Website's Usability URL:<http://blog.thewholebraingroup.com/conducting-quick-dirty-evaluation-websites-usability> [accessed 2019-08-24]
78. Thielsch MT, Blotenberg I, Jaron R. User evaluation of websites: From first impression to recommendation. Interact Comput 2014;26(1):89-102. [doi: [10.1093/iwc/iwt033](https://doi.org/10.1093/iwc/iwt033)]
79. Tung LL, Xu Y, Tan FB. Attributes of web site usability: a study of web users with the repertory grid technique. Int J Elect Commer 2009;13(4):97-126. [doi: [10.2753/jec1086-4415130405](https://doi.org/10.2753/jec1086-4415130405)]
80. Agarwal R, Venkatesh V. Assessing a firm's Web presence: a heuristic evaluation procedure for the measurement of usability. Inform Syst Res 2002;13(2):168-186. [doi: [10.1287/isre.13.2.168.84](https://doi.org/10.1287/isre.13.2.168.84)]
81. Venkatesh V, Ramesh V. Web and wireless site usability: understanding differences and modeling use. Manag Inf Syst Q 2006;30(1):181-206. [doi: [10.2307/25148723](https://doi.org/10.2307/25148723)]
82. Vaananen-Vainio-Mattila K, Wäljas M. Development of Evaluation Heuristics for Web Service User Experience. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 2009 Presented at: CHI'09; April 4-9, 2009; Boston, MA, USA p. 3679-3684. [doi: [10.1145/1520340.1520554](https://doi.org/10.1145/1520340.1520554)]

83. Wang WT, Wang B, Wei YT. Examining the Impacts of Website Complexities on User Satisfaction Based on the Task-technology Fit Model: An Experimental Research Using an Eyetracking Device. In: Proceedings of the 18th Pacific Asia Conference on Information Systems. 2014 Presented at: PACIS'14; June 18-22, 2014; Jeju Island, South Korea.
84. Yen B, Hu PJ, Wang M. Toward an analytical approach for effective web site design: a framework for modeling, evaluation and enhancement. *Elect Commer Res Appl* 2007;6(2):159-170. [doi: [10.1016/j.eierap.2006.11.004](https://doi.org/10.1016/j.eierap.2006.11.004)]
85. Yen PY, Bakken S. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. *AMIA Annu Symp Proc* 2009 Nov 14;2009:714-718 [FREE Full text] [Medline: [20351946](https://pubmed.ncbi.nlm.nih.gov/20351946/)]
86. Allison R, Hayes C, Young V, McNulty CAM. Evaluation of an educational health website on infections and antibiotics: a mixed-methods, user-centred approach in England. *JMIR Formative Research* 2019 (forthcoming).

Abbreviations

E-commerce: electronic commerce

Edited by G Eysenbach; submitted 26.04.19; peer-reviewed by C Eley, C Brown; comments to author 31.05.19; revised version received 24.06.19; accepted 18.08.19; published 19.10.19

Please cite as:

Allison R, Hayes C, McNulty CAM, Young V

A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb

JMIR Form Res 2019;3(4):e14372

URL: <http://formative.jmir.org/2019/4/e14372/>

doi: [10.2196/14372](https://doi.org/10.2196/14372)

PMID:

©Rosalie Allison, Catherine Hayes, Clodna A M McNulty, Vicki Young. Originally published in *JMIR Formative Research* (<http://formative.jmir.org>), 19.10.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <http://formative.jmir.org>, as well as this copyright and license information must be included.